

<https://helda.helsinki.fi>

Accurate genetic diagnosis of Finnish pulmonary arterial hypertension patients using oligonucleotide-selective sequencing

Vattulainen, Sanna

2015

Vattulainen , S , Aho , J , Salmenperä , P , Bruce , S , Tallila , J , Gentile , M , Sankelo , M , Laitinen , T , Koskenvuo , J W , Alastalo , T-P & Myllykangas , S 2015 , ' Accurate genetic diagnosis of Finnish pulmonary arterial hypertension patients using oligonucleotide-selective sequencing ' , Molecular Genetics & Genomic Medicine , vol. 3 , no. 4 , pp. 354-362 . <https://doi.org/10.1002/mgg3.1>

<http://hdl.handle.net/10138/233220>

<https://doi.org/10.1002/mgg3.147>

cc_by

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

ORIGINAL ARTICLE

Accurate genetic diagnosis of Finnish pulmonary arterial hypertension patients using oligonucleotide-selective sequencing

Sanna Vattulainen¹, Joonas Aho², Pertteli Salmenperä³, Siina Bruce³, Jonna Tallila³, Massimiliano Gentile³, Marja Sankelo⁴, Tarja Laitinen⁵, Juha W. Koskenvuo^{2,3,6}, Tero-Pekka Alastalo^{1,3} & Samuel Myllykangas^{3,7}

¹Pediatric Cardiology, Children's Hospital Helsinki, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland

²Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland

³Blueprint Genetics, Helsinki, Finland

⁴Department of Internal Medicine, Tampere University Hospital, Tampere, Finland

⁵Department of Pulmonary Diseases and Allergology, Turku University Hospital and University of Turku, Turku, Finland

⁶Department of Clinical Physiology and Nuclear Medicine, HUS Medical Imaging Center, Helsinki University of Central Hospital and University of Helsinki, Helsinki, Finland

⁷Department of Biochemistry and Developmental Biology, Faculty of Medicine, University of Helsinki, Helsinki, Finland

Keywords

BMPR2, genetic diagnostics, genetics, OS-Seq, pulmonary arterial hypertension.

Correspondence

Samuel Myllykangas, Department of Biochemistry and Developmental Biology, Faculty of Medicine, University of Helsinki Helsinki, Finland.

Tel: +358 50 343 9621;

E-mail: samuel.myllykangas@helsinki.fi

Funding Information

This work was supported by grants from Sigrid Jusélius Foundation, Academy of Finland, Finnish Medical Foundation, Finnish Foundation for Pediatric Research, Finnish Foundation for Cardiovascular Research, Clinical Research funding of Turku University Hospital, and Blueprint Genetics.

Received: 4 February 2015; Accepted: 24 March 2015

Molecular Genetics & Genomic Medicine
2015; 3(4): 354–362

doi: 10.1002/mgg3.147

Introduction

The role of genetic diagnostics in the evaluation of patients with Mendelian disorders is increasing rapidly in all fields of medicine. Understanding the underlying genetic causes of inherited diseases can contribute to the treatment and follow-up strategies, and may bring psy-

Abstract

The genetic basis of pulmonary arterial hypertension (PAH) among Finnish PAH patients is poorly understood. We adopted a novel-targeted next-generation sequencing (NGS) approach called Oligonucleotide-Selective Sequencing (OS-Seq) and developed a custom data analysis and interpretation pipeline to identify pathogenic base substitutions, insertions, and deletions in seven genes associated with PAH (*BMPR2*, *BMPR1B*, *ACVRL1*, *ENG*, *SMAD9*, *CAV1*, and *KCNK3*) from Finnish PAH patients. This study represents the first clinical study with OS-Seq technology on patients suffering from a rare genetic disorder. We analyzed DNA samples from 21 Finnish PAH patients, whose *BMPR2* and *ACVRL1* mutation status had been previously studied using Sanger sequencing. Our sequencing panel covered 100% of the targeted base pairs with >15× sequencing depth. Pathogenic base substitutions were identified in the *BMPR2* gene in 29% of the Finnish PAH cases. Two of the pathogenic variant-positive patients had been previously tested negative using Sanger sequencing. No clinically significant variants were identified in the six other PAH genes. Our study validates the use of targeted OS-Seq for genetic diagnostics of PAH and revealed pathogenic variants that had been previously missed using Sanger sequencing.

chological comfort for the patient. Moreover, genetic diagnosis of the proband can allow effective risk assessment of family members, rationalize follow-up strategies, and enable earlier interventions.

Pulmonary arterial hypertension (PAH) is a severe and progressive disease characterized by vascular remodeling of small pulmonary arteries resulting in an increase in

pulmonary arterial pressure, and eventually leading to right heart failure (Runo and Loyd 2003). Despite the improvement of treatments, PAH still remains a fatal disease with high mortality (Chakinala 2005). While the pathogenesis of idiopathic and hereditary forms of PAH (IPAH and HPAH) is a complex and incompletely understood, their genetic basis is well recognized. Hundreds of mutations in eight genes have been reported to associate with both IPAH and HPAH. HPAH is commonly regarded as an autosomal dominant disease (Larkin et al. 2012; Machado et al. 2009).

Mutations in the bone morphogenetic protein receptor 2 (*BMPR2*) gene, a member of the transforming growth factor beta (*TGF-β*) superfamily, are currently considered as main causes for the pathogenesis of PAH (Machado et al. 2006). Over 300 different *BMPR2* mutations have been identified in PAH patients (Machado et al. 2009, 2001). Mutations in the *BMPR2* gene have been identified in up to 75% of HPAH cases and in 25% of IPAH cases. Previously, the penetrance of PAH among *BMPR2* mutation carriers was considered low, however, a recent study based on the Varderbilt Pulmonary Hypertension Registry found an overall penetrance of 27%, (42% in females and 14% in males). The disease penetrance could be even higher as the disease manifestation may be delayed for 75 years and express phenotypic variability (Larkin et al. 2012; Machado et al. 2009; Soubrier et al. 2013).

Other genes than *BMPR2* have been implicated in rare cases of PAH. Mutations in bone morphogenetic protein receptor 1B (*BMPR1B*), activin-like kinase-type 1 (*ACVRL1*; *ALK1*), endoglin (*ENG*), smad family member 9 (*SMAD9*), caveolin 1 (*CAV1*), potassium channel subfamily K, member 3 (*KCNK3*), and *NOTCH3* have been described in the development of the disease (Austin et al. 2012; Chaouat et al. 2004; Chida et al. 2012a; Harrison et al. 2003; Li et al. 2009; Ma et al. 2013; Shintani et al. 2009). The role of these genes in the pathogenesis of IPAH and HPAH is suggested significantly smaller compared to *BMPR2*.

The genetic factors contributing to PAH in Finland, a genetic isolate in Europe, has not been comprehensively evaluated. In this study, we analyzed Finnish PAH patient isolate to detect pathogenic or likely pathogenic variants associated with PAH by utilizing the novel-targeted Oligonucleotide-Selective Sequencing (OS-Seq) technology (Myllykangas et al. 2011). We designed target-specific oligonucleotides to capture all coding exons, exon-intron boundaries, and known intronic mutations in the PAH-associated target genes. All these target genes were comprehensively analyzed using OS-Seq. These 21 Finnish PAH patients had been previously tested for mutations in the *BMPR2* and *ACVRL1* genes using Sanger sequencing (Sankelo et al. 2005). With OS-Seq we detected all previ-

ously identified pathogenic or likely pathogenic variants of *BMPR2* gene and identified two new pathogenic *BMPR2* variants that were previously missed using Sanger sequencing. Our results validate the use of OS-Seq panel for diagnostics of PAH.

Methods

Patient samples

Blood samples for genetic studies were obtained from Finnish patients diagnosed with PAH under the approval from the Ethical Committees of the five University Hospitals and by the Ministry of Social Affairs and Health of Finland. All participants provided the written informed consent for genetic analysis as previously reported (Sankelo et al. 2005). Genomic DNA was extracted from the blood samples (Sankelo et al. 2005) and DNA samples of 21 Finnish iPAH (*N* = 18) and HPAH (*N* = 3) patients were selected for the study.

PAH panel

Seven genes were selected for the PAH panel based on literature search in 2012–2013 (*BMPR2*, *BMPR1B*, *ACVRL1*, *ENG*, *SMAD9*, *CAV1*, and exon 2 of *KCNK3*) and all validated information about the mutations were incorporated into the bioinformatics pipeline (Table 1 and Fig. 1). Noteworthy, exon 1 of *KCNK3* was omitted in the analysis because it is difficult to sequence due to extremely high GC content and because all mutations in *KCNK3* that have been linked with PAH are located in exon 2 (Ma et al. 2013).

Sequencing

The quality of genomic DNA samples was evaluated using gel electrophoresis and DNA concentration was measured using Qubit (Life Technologies, Grand Island, NY). Geno-

Table 1. Genes selected for diagnostics of PAH.

Gene symbol	Gene name	Reference
<i>BMPR2</i>	Bone morphogenic protein receptor 2	Machado et al. (2006)
<i>BMPR1B</i>	Bone morphogenic protein receptor 1B	Chida et al. (2012a, b)
<i>ACVRL1</i>	Activin A receptor type II like-1	Harrison et al. (2003)
<i>ENG</i>	Endoglin	Chaouat et al. (2004)
<i>SMAD9</i>	Smad family member 9	Shintani et al. (2009)
<i>CAV1</i>	Caveolin 1	Austin et al. (2012)
<i>KCNK3</i>	Potassium channel, subfamily K, member 3	Ma et al. (2013)

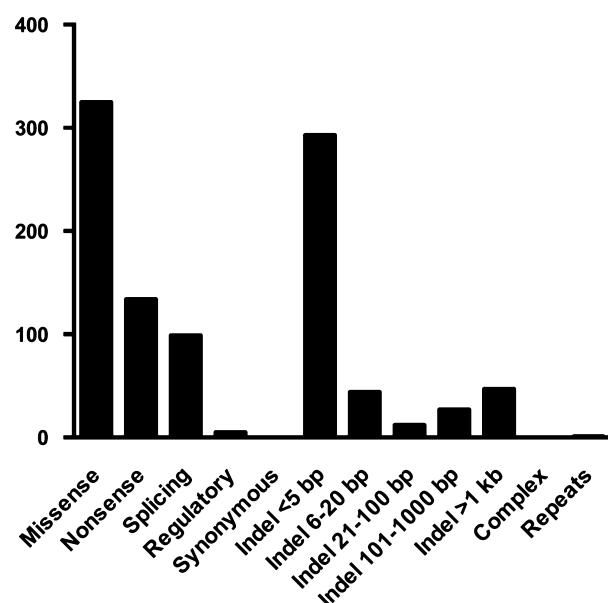


Figure 1. Distribution of mutations in genes selected for diagnostics of PAH. Mutations in seven PAH-associated genes were collected from the literature.

mic DNA samples were fragmented using Fragmentase (New England Biolabs, Ipswich, MA) and sequencing libraries were prepared using NEBNext Ultra (New England Biolabs) and PCR amplified. Target-specific oligonucleotides were designed to capture all coding exons, exon-intron boundaries, and known intronic mutations in the PAH-associated target genes and obtained from Integrated DNA Technologies (Leuven, Belgium). Exon definitions for target genes were derived from the CCDS. The targeting and sequencing of DNA was automated using the MiSeq Desktop Sequencer (Illumina, San Diego, CA). Pathogenic and likely pathogenic mutations found in the patient samples were confirmed using Sanger sequencing.

NGS data analysis

Primary data analysis was carried out by Illumina's MiSeq Reporter software using the generateFASTQ workflow, that generates a single FASTQ file with base calls and Phred quality scores for R1 and R2 reads, separately. The data were transferred to a dedicated pipeline server for downstream analysis, which started with extraction of data for individual samples using a proprietary demultiplexing algorithm that analyses the 9 bp long index adapter sequence of each read to determine sample identity. Sequence reads were then trimmed to exclude adapter and capture oligo sequences using a proprietary algorithm that estimates the size of captured target DNA based on the genomic coordinates of read mate-pairs. Low-quality data were identified and removed using the Prinseq

package (Schmieder and Edwards 2011) to exclude reads with more than 10% of unclassified nucleotides, as well as trimming nucleotides with Phred scores inferior to 20 from starting from the 3'-end. The cleaned-up sequence reads were subsequently aligned to the human reference sequence assembly GRCh37 using the Burrows-Wheeler Aligner (Li and Durbin 2009) in single-end mode with default parameters. Variant calling was performed using the Genome Analysis Tool Kit (GATK) (McKenna et al. 2010) following the guidelines outlined in the document "Best Practice Variant Detection with the GATK (v1.x)" issued by the developers.

Variant interpretation

An automated bioinformatics pipeline was applied to identify pathogenic base substitutions, insertions and deletions (INDELs) in PAH-associated genes. Mutations that have been shown to be associated with PAH were collected from original research publications and existing mutation databases and queried using patients' variants. Population frequencies for detected variants were derived from the Exome Variant Server (NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (<http://evs.gs.washington.edu/EVS>)) and the Exome Aggregation Consortium (ExAC), Cambridge, MA (<http://exac.broadinstitute.org>). PolyPhen, SIFT (Adzhubei et al. 2010; Ng and Henikoff 2001, 2002), and Mutation Taster (Schwarz et al. 2014) tools were used to predict the effects of variants on the protein structure and functions. Variants were classified as pathogenic or likely pathogenic when previous studies had identified them in other PAH patients. Furthermore, pathogenic gene variants associated with PAH are considered to be extremely rare or absent in control populations and often considered deleterious by in silico prediction.

Results

Here, we present the first clinical study on Finnish PAH patients using the novel-targeted OS-Seq technology. We confirmed pathogenic or likely pathogenic variants of *BMPR2* in four patients and detected two new pathogenic variants of *BMPR2* gene that were previously missed using Sanger sequencing (Sankelo et al. 2005). Four out of six patients were carrying pathogenic gene variants that were predicted to cause a truncated *BMPR2* protein: the point mutations c.994C>T (p.Arg332X) in the kinase domain and c.1750C>T (p.Arg584X) in the C-terminal cytoplasmic domain, and a deletion causing a frameshift mutation (c.1376_1377delGA, p.Arg459ThrfsX11) in the kinase domain in two patients. In addition, two of the patients had missense point mutations in *BMPR2*: c.1472G>A (p.Arg491Gln) located in a kinase domain and

c.2696G>C (p.Arg899Pro) located in C-terminal cytoplasmic domain (Table 2). Both of these variants featuring missense mutations were classified as deleterious in all in silico variant effect prediction analyses. In addition to these five pathogenic variants, three other variants were detected in *BMPR2*. Polymorphisms were observed in two patients and one short GGC-insertion in the promoter was identified in three patients (Table 3). The clinical significance of the identified GGC-insertion is unknown, since the frequency of this variant is unknown in control populations, as the region is not covered in these databases. Altogether, our OS-Seq panel revealed six (29%) PAH patients carrying a pathogenic, or likely pathogenic, variant of *BMPR2* gene (Table 2).

In addition, seven variants were detected in *ACVRL1* and *ENG* (Table 3). One patient carried a gene variant featuring a novel missense mutation c.536A>C (p.Asp179Ala) in a serine and threonine residue-rich region in the *ACVRL1* gene. This *ACVRL1* variant was predicted to be benign by in silico analyses. This patient also tested positive for a pathogenic variant featuring the frameshift mutation in *BMPR2* (Table 3), and therefore the *ACVRL1* variant was classified as likely benign. The variants identified in the *ENG* gene were not considered deleterious (Table 3).

We evaluated the average sequencing depth of the PAH panel for the DNA samples of the 21 Finnish PAH patients by measuring the number of overlapping

Table 2. Mutations in *BMPR2* in Finnish iPAH and hPAH patients.

Patient	1	2	3	4	5	6
Sex	F	F	M	M	F	F
Age of diagnosis	31	52	35	49	19	47
Family history	No	No	Yes	No	No	No
Exon number	8	10	11	12	10	12
Nucleotide change	c.994C>T	c.1376_1377delGA	c.1472G>A	c.2696G>C	c.1376_1377delGA	c.1750C>T
Amino acid change	p.Arg332X	p.Arg459ThrfsX11	p.Arg491Gln	p.Arg899Pro	p.Arg459ThrfsX11	p.Arg584X
Mutation type	Stop	Frameshift	Missense	Missense	Frameshift	Stop
Protein domain	Kinase domain	Kinase domain	Kinase domain	C-terminal cytoplasmic domain	Kinase domain	C-terminal cytoplasmic domain
Detected previously by Sanger Sequencing	Yes	No	Yes	Yes	Yes	No
Mutation in other genes	None	<i>ACVRL1</i> *	None	None	None	None

Likely benign (c.536A>C, p.Asp179Ala). Mutation nomenclature is based on GenBank accession NM_001204.6 with nucleotide on being the first nucleotide of the translation initiation codon.

* Refers to * Likely benign (c.536A<C, p.Asp179Ala).

Table 3. Variants identified in Finnish PAH patients.

Gene	Variant ID	Consequence	Clinical significance	Variant prediction	MAF ¹	Number of patients
ACVRL1	CM033589	Missense	Likely benign	Tolerated	0	1
BMPR2	rs137852751	Stop	Pathogenic	Disease-causing	0	1
BMPR2	–	Frameshift	Pathogenic	Disease-causing	0	2
BMPR2	rs137852752	Missense	Pathogenic	Disease-causing	0.000008134	1
BMPR2	rs137852749	Missense	Pathogenic	Disease-causing	0	1
BMPR2	CM010166, COSM209591	Stop	Likely pathogenic	Disease-causing	0	1
BMPR2	rs1061157	Synonymous	Benign	Neutral	0.1201	7
BMPR2	rs375624016	5'UTR-insertion	Unknown	Unknown	0	3
BMPR2	rs55722784	Synonymous	Benign	Neutral	0.01371	2
ENG	rs3739817	Synonymous	Benign	Neutral	0.09142	4
ENG	rs34828244	Synonymous	Benign	Neutral	0.01042	2
ENG	rs36092484	Synonymous	Benign	Neutral	0.01801	2
ENG	rs16930129 (rs11545664)	Synonymous	Benign	Neutral	0.0983	4
ENG	rs35400405	Missense	Benign	Tolerated	0.04790	2
ENG	rs7847860	Intron	Benign	Neutral	0.06638	3

¹Minor allele frequencies were collected from the ExAC-project data.

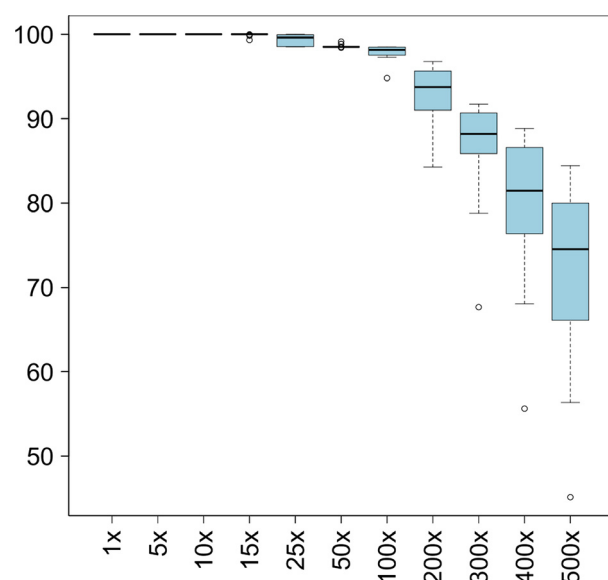


Figure 2. Percentage of target bases covered at different sequencing depths in 21 Finnish PAH patient samples. Results from the sequencing coverage analyses are shown as percentage of targeted bases with $>1\times$, $>5\times$, $>10\times$, $>15\times$, $>25\times$, $>50\times$, $>100\times$, and until $>500\times$ depth. To generate highly confident genotype calls, $>15\times$ coverage is considered to be enough.

sequencing reads at each nucleotide position in the 12,638 base target region. The median sequencing depth was $791\times$ and coverage of nucleotides with $>15\times$ sequencing depth was 100%. To demonstrate the sequencing coverage performance of the PAH panel we calculated the percentage of nucleotides exceeding specific sequencing depth threshold ($1\times$ – $500\times$) (Fig. 2). To demonstrate further the efficiency of targeted sequencing, the sequencing depth was also evaluated for individual patient samples (Fig. 3) showing that 100% of the target regions were covered with $>15\times$ sequencing depth. Furthermore, we showed that the sequencing depth in the coding exons of the *BMP2* gene in six PAH patient samples were sequenced in at high depth, as the median coverage was $978\times$ and the coverage of nucleotides with $>15\times$ coverage was 100% (Fig. 4).

Discussion

Diagnostics of inherited diseases has entered a new era, in which sophisticated sequencing and bioinformatics reveal pathogenic mutations efficiently. There is estimated to be more than 7000 disorders with Mendelian inheritance and for many of these diseases the molecular genetic mechanism has been discovered. Although individual inherited diseases are rare, together they impose a significant clinical and economical challenge (Costa et al. 1985).

The increasing realization of genetic underpinnings of inherited diseases has led to diagnostic challenges. Obtaining a complete genetic view of the disease requires high-quality sequencing of large number of genes and genomic regions with explicit clinical relevance and is beyond the scope and capacity of Sanger sequencing applications. At the same time, treatment of inherited diseases in the clinic presumes faster turnaround time and cost-effectiveness from genetic testing. Therefore, long odysseys of mutation exclusion by queuing Sanger sequencing experiments are already outdated in the clinic. Novel technologies have emerged to fulfill the growing clinical demand and to overcome the challenges associated with genetic testing of inherited diseases.

Targeted next-generation sequencing (NGS) is a practical solution for diagnostic-grade and comprehensive analysis of the clinically relevant genomic regions. While sequencing whole genomes or exomes are powerful tools for discovery and research, they fall short meeting the quality requirements of genetic diagnostics. Typical genome with $30\times$ coverage or exome even with $100\times$ coverage contains gaps and regions where sequence information remains unreliable (Meynert et al. 2013; Rehm et al. 2013; Sims et al. 2014). In order to reach consummate clinical genetic interpretation and solid diagnosis, those regions where data are narrow or missing need to be patched with alternative approaches, such as Sanger Sequencing. Necessity of gap-filling makes application of whole-genome sequencing (WGS) and whole-exome sequencing (WES) for genetic diagnostics uneconomical and impractical. Targeted NGS approaches enable cost-efficient, reliable, and high-depth sequencing of clinically relevant genomic regions with complete coverage.

To demonstrate the utility of NGS in the diagnostics of inherited diseases, we have developed an OS-Seq sequencing panel to accurately detect mutations in seven genes that have been implicated in the pathogenesis of PAH. We collected over 900 mutations from literature and existing databases to support the interpretation of the patients' variants. Over 90% of the collected mutations were base substitutions causing missense, nonsense, and splice site mutations or small indels of up to 20 bp that can be identified with NGS (Fig. 1).

The overall performance of our test was high. Sequencing depth and coverage are the main indicators of the performance of NGS. Sequencing depth refers to the number of sequencing reads that pileup at specific nucleotide and is directly associated with the confidence of the genotype call. The average (measured as median) sequencing depth was 791, which illustrates the confidence in calling the genotypes. Sequencing coverage defines the breadth of the genomic area, which has sequencing depth exceeding a specific threshold.

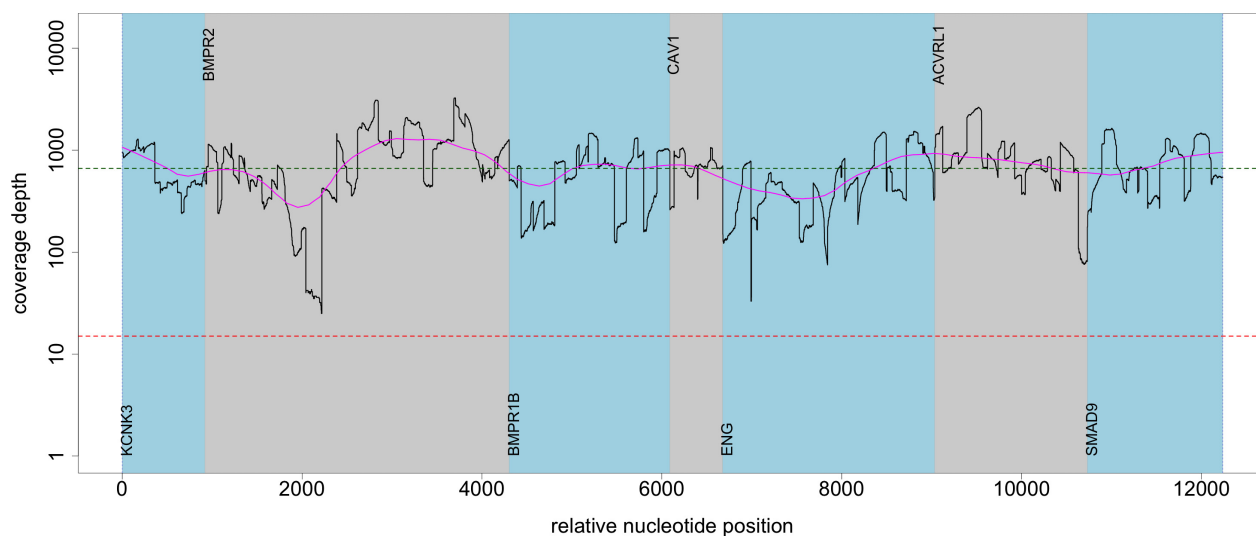


Figure 3. Sequencing coverage profile of an individual patient sample. The coverage profile of an individual patient shows how OS-Seq PAH panels with median sequencing depth of 662 and covered 100% target regions with $>15\times$ coverage. Coverage depth in y-axis shows the average of reads the certain nucleotide is sequenced. Relative nucleotide position is shown in x-axis.

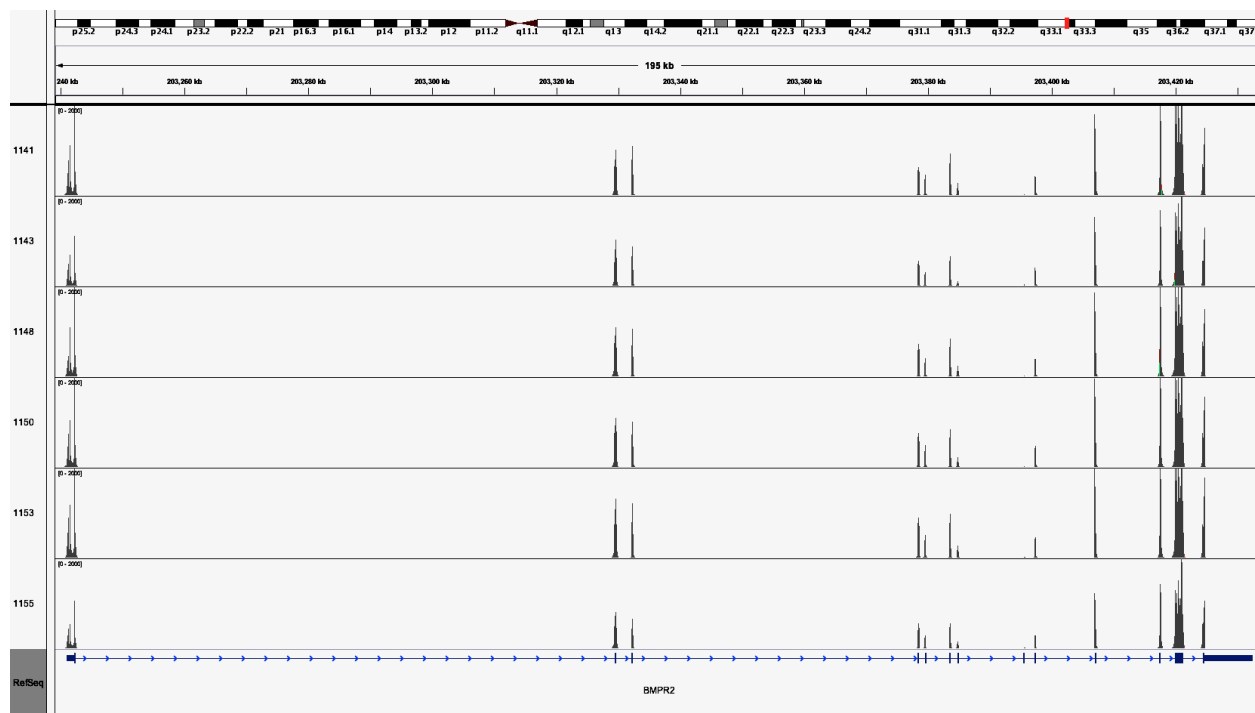


Figure 4. Coverage plot of coding exons of the *BMPR2* gene. The sequencing depth in the coding exons of the *BMPR2* gene in six PAH patient samples were sequenced in at high depths. The median coverage was 978 and coverage of nucleotides with $>15\times$ sequencing depth was 100%.

Sequencing coverage is related to sensitivity of detecting variants in the target regions (Sims et al. 2014). We regarded target regions being covered when exceeding a $15\times$ sequencing depth. It has been demonstrated that a

sequencing depth of at least $15\times$ is appropriate for making confident calls, as calling heterozygous variants rarely requires more than 13 overlapping sequence reads (Meynert et al. 2013; Sims et al. 2014). Sequencing coverage at

>15× depth for our PAH panel was 100% indicating a high sensitivity to detect variants in the target regions.

We analyzed 21 Finnish PAH patients using our novel, targeted OS-Seq panel for PAH. Although the study cohort was small it was able to confirm relatively high incidence (29%) of pathogenic or likely pathogenic variants in *BMPR2* gene among Finnish PAH patients (Table 2). In this cohort, the analysis of other PAH-associated genes did not reveal any other genetic variants considered significant (Table 3), further pinpointing the pivotal role that *BMPR2* has in the pathogenesis of PAH. Importantly, with our OS-Seq approach, we identified all previously detected *BMPR2* mutations and identified two individuals with pathogenic or likely pathogenic *BMPR2* gene variant that were previously missed by Sanger sequencing. In the Sanger method, nonspecific binding of the primers and the formation of DNA secondary structures may cause sequencing errors (Hert et al. 2008; Sanger et al. 1977). There is also increased risk for human error when interpreting the raw sequencing results due to ambiguities in the capillary electrophoresis readouts. These factors could have caused the failures to identify the two *BMPR2* mutations in the primary study (Sankelo et al. 2005).

ACVRL1 or *ENG* mutations are often found in PAH patients with family history of hereditary hemorrhagic telangiectasia (HHT) (Harrison et al. 2003). As none of the studied Finnish PAH patients manifested clear HHT-disease, it was not surprising that the analysis of *ACVRL1* and *ENG* did not identify any pathogenic variants (Table 3). The patient with the variant in the *ACVRL1* gene (c.536A>C, p.Asp179Ala) also carried a pathogenic deletion in the *BMPR2* gene (c.1376_1377delGA, p.Arg459ThrfsX11) (Table 2). Interestingly, this pathogenic *BMPR2* variant resulting a frameshift was missed in previous study (Sankelo et al. 2005) using Sanger sequencing and the *ACVRL1* mutation was considered a potential cause for PAH although the patient lacked clear symptoms associated with HHT. As the patient is carrying a *BMPR2* variant resulting a deletion and is considered to be pathogenic, we suggest that it is unlikely that this *ACVRL1* variant is an independent disease-causing mutation. The identified *ENG* gene variant, featuring a missense mutation (c.14C>T, p.Thr5Met; rs35400405), from two patients is interesting as it leads to the replacement of threonine by methionine resulting a suppression of *ENG* expression (Bourdeau et al. 2001). Despite its direct effects on protein (Bourdeau et al. 2001) it is predicted to be non-disease causing as the allele is common in general population (MAF 0.04790) (Table 3).

Characterization of mutations in PAH enables establishing estimates on prognosis. *BMPR2* mutations are associated with an earlier age of disease onset, regardless

of familial history, and with more severe disease phenotype. Furthermore, patients with *BMPR2* mutations are less likely to respond to vasodilators than mutation-negative patients (Austin et al. 2009; Ma and Chung 2014). PAH patients carrying mutations in *ACVRL1* gene show more rapid disease progression than patients with *BMPR2* mutations, despite responding to vasodilators at the time of diagnosis (Girerd et al. 2010). In addition, patients identified with *BMPR2* and *ACVRL1* mutations have a shorter survival time or earlier need for lung transplantation (Austin et al. 2009; Chida et al. 2012b; Girerd et al. 2010; Ma and Chung 2014). The study of Austin and coworkers in 2009 reported that patients with missense mutations in *BMPR2* gene are diagnosed in much younger age and seem to have a significantly shorter survival time compared to patients with truncating mutations (Austin et al. 2009). In 2012, Liu with colleagues reported male *BMPR2* mutation carriers had a worse prognosis compared to female mutation carriers (Liu et al. 2012). Pediatric PAH patients, with or without *BMPR2* mutation, show distinct response to vasodilators than adults, and are more likely to associate with other genetic syndromes (Ma and Chung 2014).

Utilization of novel targeted OS-Seq approach increases the diagnostic efficacy by offering better quality and faster turnaround time. As shown in this study, OS-Seq technology proved to be a practical and effective tool for genetic profiling of Finnish patients with PAH. As genetic testing is increasingly applied in clinical practice, it is important to acknowledge the requirements of quality and performance of available genetic tests. Diagnostic-grade deep sequencing with 100% covered target base pairs is becoming a standard requirement in today's clinical genetic testing. The high incidence of disease mutations in patients with IPAH and HPAH and elevated estimates of disease penetrance are supporting the utilization of genetic testing as routine procedure in the evaluation of patients with IPAH or HPAH. As in other hereditary cardiovascular diseases genetic diagnosis can significantly rationalize the risk stratification and follow-up strategies in the family and can have impact in estimating index patient's prognosis. New sequencing strategies and bioinformatics tools are enabling diagnostics with faster turnaround times and decreased cost.

Acknowledgments

The authors thank the Finnish PAH patients and their families for their contribution into this study, M.Sc. Janica Djupsjöbacka (Blueprint Genetics) for excellent technical assistance, and the NHLBI GO Exome Sequencing Project and its ongoing studies which produced and

provided exome variant calls for comparison: the Lung GO Sequencing Project (HL-102923), the WHI Sequencing Project (HL-102924), the Broad GO Sequencing Project (HL-102925), the Seattle GO Sequencing Project (HL-102926) and the Heart GO Sequencing Project (HL-103010), and The Exome Aggregation Consortium (ExAC) and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at <http://exac.broadinstitute.org/about>.

Conflict of Interest

SV, JA, MS, and TL have no conflicts of interest to declare. PS, SB, JT, and MG are employees of Blueprint Genetics, which develops genetic diagnostic applications for clinical diagnostics and also offers diagnostics for PAH. JWK, TPA, and SM are cofounders of Blueprint Genetics.

References

- Adzhubei, I. A., S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova, P. Bork, et al. 2010. A method and server for predicting damaging missense mutations. *Nat. Methods* 7:248–249.
- Austin, E. D., J. A. Phillips, J. D. Cogan, R. Hamid, C. Yu, K. C. Stanton, et al. 2009. Truncating and missense BMPR2 mutations differentially affect the severity of heritable pulmonary arterial hypertension. *Respir. Res.* 10:87.
- Austin, E. D., L. Ma, C. LeDuc, E. Berman Rosenzweig, A. Borczuk, J. A. 3rd Phillips, et al. 2012. Whole exome sequencing to identify a novel gene (caveolin-1) associated with human pulmonary arterial hypertension. *Circ. Cardiovasc. Genet.* 5:336–343.
- Bourdeau, A., M. E. Faughnan, M. L. McDonald, A. D. Paterson, I. R. Wanless, and M. Letarte. 2001. Potential role of modifier genes influencing transforming growth factor- β 1 levels in the development of vascular defects in endoglin heterozygous mice with hereditary hemorrhagic telangiectasia. *Am. J. Pathol.* 158:2011–2020.
- Chakinala, M. M. 2005. Changing the prognosis of pulmonary arterial hypertension: impact of medical therapy. *Semin. Respir. Crit. Care Med.* 26:409–416.
- Chaouat, A., F. Coulet, C. Favre, G. Simonneau, E. Weitzenblum, F. Soubrier, et al. 2004. Endoglin germline mutation in a patient with hereditary haemorrhagic telangiectasia and dexfenfluramine associated pulmonary arterial hypertension. *Thorax* 59:446–448.
- Chida, A., M. Shintani, T. Nakayama, Y. Furutani, E. Hayama, K. Inai, et al. 2012a. Missense mutations of the BMPR1B (ALK6) gene in childhood idiopathic pulmonary arterial hypertension. *Circ. J.* 76:1501–1508.
- Chida, A., M. Shintani, H. Yagi, M. Fujiwara, Y. Kojima, H. Sato, et al. 2012b. Outcomes of childhood pulmonary arterial hypertension in BMPR2 and ALK1 mutation carriers. *Am. J. Cardiol.* 110:586–593.
- Costa, T., C. R. Scriver, and B. Childs. 1985. The effect of Mendelian disease on human health: a measurement. *Am. J. Med. Genet.* 21:231–242.
- Girerd, B., D. Montani, F. Coulet, B. Sztrymf, A. Yaici, X. Jais, et al. 2010. Clinical outcomes of pulmonary arterial hypertension in patients carrying an ACVRL1 (ALK1) mutation. *Am. J. Respir. Crit. Care Med.* 181:851–861.
- Harrison, R. E., J. A. Flanagan, M. Sankelo, S. A. Abdalla, J. Rowell, R. D. Machado, et al. 2003. Molecular and functional analysis identifies ALK-1 as the predominant cause of pulmonary hypertension related to hereditary haemorrhagic telangiectasia. *J. Med. Genet.* 40:865–871.
- Hert, D. G., C. P. Fredlake, and A. E. Barron. 2008. Advantages and limitations of next-generation sequencing technologies: a comparison of electrophoresis and non-electrophoresis methods. *Electrophoresis* 29:4618–4626.
- Larkin, E. K., J. H. Newman, E. D. Austin, A. R. Hemnes, L. Wheeler, I. M. Robbins, et al. 2012. Longitudinal analysis casts doubt on the presence of genetic anticipation in heritable pulmonary arterial hypertension. *Am. J. Respir. Crit. Care Med.* 186:892–896.
- Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760.
- Li, X., X. Zhang, R. Leathers, A. Makino, C. Huang, P. Parsa, et al. 2009. Notch3 signaling promotes the development of pulmonary arterial hypertension. *Nat. Med.* 15:1289–1297.
- Liu, D., W. H. Wu, Y. M. Mao, P. Yuan, R. Zhang, F. L. Ju, et al. 2012. BMPR2 mutations influence phenotype more obviously in male patients with pulmonary arterial hypertension. *Circ. Cardiovasc. Genet.* 5:511–518.
- Ma, L., and W. K. Chung. 2014. The genetic basis of pulmonary arterial hypertension. *Hum. Genet.* 133:471–479.
- Ma, L., D. Roman-Campos, E. D. Austin, M. Eyries, K. S. Sampson, F. Soubrier, et al. 2013. A novel channelopathy in pulmonary arterial hypertension. *N. Engl. J. Med.* 369:351–361.
- Machado, R. D., M. W. Pauciuolo, J. R. Thomson, K. B. Lane, N. V. Morgan, L. Wheeler, et al. 2001. BMPR2 haploinsufficiency as the inherited molecular mechanism for primary pulmonary hypertension. *Am. J. Hum. Genet.* 68:92–102.
- Machado, R. D., M. A. Aldred, V. James, R. E. Harrison, B. Patel, E. C. Schwalbe, et al., et al. 2006. Mutations of the TGF- β type II receptor BMPR2 in pulmonary arterial hypertension. *Hum. Mutat.* 27:121–132.
- Machado, R. D., O. Eickelberg, C. G. Elliott, M. W. Geraci, M. Hanaoka, J. E. Loyd, et al. 2009. Genetics and genomics of pulmonary arterial hypertension. *J. Am. Coll. Cardiol.* 54(1 Suppl.):S32–S42.

- McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, et al. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20:1297–1303.
- Meynert, A. M., L. S. Bicknell, M. E. Hurles, A. P. Jackson, and M. S. Taylor. 2013. Quantifying single nucleotide variant detection sensitivity in exome sequencing. *BMC Bioinform.* 14:195.
- Myllykangas, S., J. D. Buenrostro, G. Natsoulis, J. M. Bell, and H. P. Ji. 2011. Efficient targeted resequencing of human germline and cancer genomes by oligonucleotide-selective sequencing. *Nat. Biotechnol.* 29:1024–1027.
- Ng, P. C., and S. Henikoff. 2001. Predicting deleterious amino acid substitutions. *Genome Res.* 11:863–874.
- Ng, P. C., and S. Henikoff. 2002. Accounting for human polymorphisms predicted to affect protein function. *Genome Res.* 12:436–446.
- Rehm, H. L., S. J. Bale, P. Bayrak-Toydemir, J. S. Berg, K. K. Brown, J. L. Deignan, et al. 2013. ACMG clinical laboratory standards for next-generation sequencing. *Genet. Med.* 15:733–747.
- Runo, J. R., and J. E. Loyd. 2003. Primary pulmonary hypertension. *Lancet* 361:1533–1544.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463–5467.
- Sankelo, M., J. A. Flanagan, R. Machado, R. Harrison, N. Rudarakanchana, N. Morrell, et al. 2005. BMPR2 mutations have short lifetime expectancy in primary pulmonary hypertension. *Hum. Mutat.* 26:119–124.
- Schmieder, R., and R. Edwards. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864.
- Schwarz, J. M., D. N. Cooper, M. Schuelke, and D. Seelow. 2014. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* 11:361–362.
- Shintani, M., H. Yagi, T. Nakayama, T. Saji, and R. Matsuoka. 2009. A new nonsense mutation of SMAD8 associated with pulmonary arterial hypertension. *J. Med. Genet.* 46:331–337.
- Sims, D., I. Sudbery, N. E. Illott, A. Heger, and C. P. Ponting. 2014. Sequencing depth and coverage: key considerations in genomic analyses. *Nat. Rev. Genet.* 15:121–132.
- Soubrier, F., W. K. Chung, R. Machado, E. Grunig, M. Aldred, M. Geraci, et al. 2013. Genetics and genomics of pulmonary arterial hypertension. *J. Am. Coll. Cardiol.* 62(25 Suppl.): D13–D21.